

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 1999, 01:30:23 ; Search time 366.45 Seconds  
(without alignments)  
3475.528 Million cell updates/sec

Title: US-09-205-015-3  
Perfect score: 356  
Sequence: 1.tcgaccctctggaacctatc.....atctggagctgaagaattc 356

Scoring table: IDENTITY\_NUC

Searched: 808301 seqs, 1788773984 residues

Database : GenEmbl.\*

- 1: gb\_ba1.\*
- 2: gb\_ba2.\*
- 3: gb\_in.\*
- 4: gb\_ov.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl1.\*
- 9: gb\_pl2.\*
- 10: gb\_pr1.\*
- 11: gb\_pr2.\*
- 12: gb\_pr3.\*
- 13: gb\_ro.\*
- 14: gb\_st.\*
- 15: gb\_sy.\*
- 16: gb\_un.\*
- 17: gb\_v1.\*
- 18: gb\_vtg.\*
- 19: em\_ba.\*
- 20: em\_fun.\*
- 21: em\_hum1.\*
- 22: em\_hum2.\*
- 23: em\_in.\*
- 24: em\_om.\*
- 25: em\_or.\*
- 26: em\_ov.\*
- 27: em\_pat.\*
- 28: em\_ph.\*
- 29: em\_pl.\*
- 30: em\_ro.\*
- 31: em\_sy.\*
- 32: em\_un.\*
- 33: em\_v1.\*
- 34: em\_vtg.\*
- 35: em\_sts.\*
- 36: gb\_ba1.\*
- 37: gb\_ba2.\*
- 38: gb\_pl1.\*
- 39: gb\_pl2.\*
- 40: gb\_pr1.\*
- 41: gb\_pr2.\*
- 42: gb\_pr3.\*
- 43: gb\_sts.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	354.4	99.6	356	10	S49899
2	354.4	99.6	19226	11	HS6G4
3	354.4	99.6	356	40	S49899
4	354.4	99.6	19226	41	HS6G4
5	102.2	28.7	277	13	S78508
6	102.2	28.7	330	13	MMU08220
7	37.8	10.6	7218	6	I66494
8	36.6	10.3	136019	18	HS327J16
9	35.8	10.1	3905	11	HSB9HB1
10	35.8	10.1	3905	41	HSB9HB1
11	33.6	9.4	2844	6	I08667
12	33.6	9.4	1858	10	HSPAI19
13	33.6	9.4	1962	10	HSPAI1
14	33.6	9.4	2876	10	HUMPA1
15	33.6	9.4	17509	10	HUMPA1A
16	33.6	9.4	2937	10	HUMPA1B
17	33.6	9.4	134641	18	AC004876
18	33.6	9.4	1858	40	HSPAI19
19	33.6	9.4	1962	40	HSPAI1
20	33.6	9.4	2876	40	HUMPA1
21	33.6	9.4	17509	40	HUMPA1A
22	33.6	9.4	2937	40	HUMPA1B
23	33.4	9.4	174311	3	AC004758
24	33.4	9.4	89818	11	AC002126
25	33.4	9.4	34877	11	CH19R27740
26	33.4	9.4	170891	18	AC002118
27	33.4	9.4	89818	41	AC002126
28	33.4	9.4	34877	41	CH19R27740
29	33	9.3	1149	10	HSPKCG5
30	33	9.3	1149	40	HSPKCG5
31	32.8	9.2	325	43	HSC003ZA5
32	32.6	9.2	1873	6	AR001351
33	32.6	9.2	1873	6	AR002247
34	32.6	9.2	1873	6	I15560
35	32.6	9.2	1873	6	I56648
36	32.6	9.2	1873	6	I58411
37	32.6	9.2	1873	6	I58687
38	32.6	9.2	1873	6	I60517
39	32.6	9.2	1873	6	I68197
40	32.6	9.2	1873	6	I79527
41	32.6	9.2	1873	6	I83789
42	32.6	9.2	1873	6	I96042
43	32.6	9.2	2351	10	HSRNCAPGB
44	32.6	9.2	112756	11	AF030876
45	32.6	9.2	2351	40	HSRNCAPGB

ALIGNMENTS

RESULT	1	S49899	356 bp	DNA	PRI	10-JUL-1992
LOCUS	S49899	alpha-globin gene cluster: [5' region, major regulatory element]				
DEFINITION	S49899	[human, Genomic, 356 nt].				
ACCESSION	S49899	g233777				
NID	S49899	human.				
KEYWORDS	S49899	human.				
SOURCE	S49899	human.				
ORGANISM	S49899	human.				
REFERENCE	S49899	1 (bases 1 to 356)				
AUTHORS	S49899	Jarman, A.P., Wood, W.G., Sharpe, J.A., Gourdon, G., Ayyub, H. and Higgins, D.R.				
TITLE	S49899	Characterization of the major regulatory element upstream of the human alpha-globin gene cluster				
JOURNAL	S49899	Mol. Cell. Biol. 11 (9), 4679-4689 (1991)				
MEDLINE	S49899	91342671				
REMARK	S49899	GenBank staff at the National Library of Medicine created this entry [NCBI gbbbsq 49899] from the original journal article. This sequence comes from fig 5.				
COMMENT	S49899	Region: alpha-globin gene cluster.				

```

FEATURES
  source      Location/Qualifiers
  1..356      /organism="Homo sapiens"
              /db_xref="taxon:9606"
BASE COUNT   89 a  89 c  108 g  70 t
ORIGIN
Query Match      99.6%; Score 354.4; DB 10; Length 356;
Best Local Similarity 99.7%; Pred. No. 6.9e-103;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tcgacctctggaacctatcagggaccacagtcagccaggcagcacatctgcccagcc 60
Db 1 TCAGCCCTCTGGAACCTATCAGGGACCCACAGTCAGCCAGGCAAGCACATCTGCCCAAGCC 60

QY 61 aagggtggagggatgcagctgtgggggtctgtgaaacacttgaggagcagataactgg 120
Db 61 AAGGGTGGAGGATGCAGCTGTGGGGGTCTGTGAAACACTTGGAGGAGCAGATAACTGG 120

QY 121 gcaaccatgactcagtcgtcttgaggccacagagactcttgatcatcctctgtgggggt 180
Db 121 GCCAACCATGACTCAGTCTCTGGAGGCCAACAGGACTGCTGACTCATCTGTGGGGGT 180

QY 181 ggagtggaacaggaaagggtgaatggtactctgattacaacctctgtgtgcct 240
Db 181 GGAGGTGGAGCAAGGAAAGGGGTGAATGTTACTCTGATTACAACTCTGTGTGCTGCT 240

QY 241 cccctctctgttattctgagagggaagccatgcccaaggtttcacagccagacttcag 300
Db 241 CCCCTCTCTGTTTACTGAGAGGGAGGCCATGCCCAAGGTTCACAGCCAGGCTTCAG 300

QY 301 gggcaagcctgaccacagacagtaatacgtttcttcctctgagctgagtgagaaattc 356
Db 301 GGGCAAGCCTGACCACAGACAGTAATACTGTTCTTCATCTGGAGCTGAAGAAATTC 356

RESULT 2
HSGG4      19226 bp  DNA  PRI  19-MAR-1997
LOCUS      Human DNA sequence from cosmid GG4 from a contig from the tip of
DEFINITION the short arm of chromosome 16, spanning 2Mb of 16p13.3.
ACCESSION  Z84722
NID        Q1817579
KEYWORDS   16p13.3.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 19226)
AUTHORS    Flint,J. and Higgs,D.R.
TITLE      Direct Submission
JOURNAL    Submitted (28-JAN-1997) Sanger Centre, 970124/Hinxton,
            Cambridgeshire, CB10 1RQ, UK. E-mail enquires: humpub@sanger.ac.uk
COMMENT    IMPORTANT: This sequence is not the entire insert of clone GG4.
            This clone was sequenced at the Institute of Molecular Medicine.
            The true left end of clone GG4 is at 1 in this sequence. The true
            right end of clone GG4 is at 456.
            The true left end of clone PX94 is at 19090.
            GG4 is from a 280kb clone contig extending from the telomere of
            16p.
            Higgs D.R., Flint J. unpublished. MRC Molecular Haematology Unit,
            Institute of Molecular Medicine, Oxford.
            GG4 came from the Los Alamos, flow sorted human Chromosome 16
            library.

FEATURES
  source      Location/Qualifiers
  1..19226    /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="16"
              /map="16p13.3"
              /clone="GG4"
              601..894
repeat_region

```

```

Query Match      99.6%; Score 354.4; DB 11; Length 19226;
Best Local Similarity 99.7%; Pred. No. 9.9e-103;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tcgacctctggaacctatcagggaccacagtcagccaggcagcacatctgcccagcc 60
|||||

```















```

      exon      /note="PAI-1, intron D"
                /number=5
      intron     /note="PAI-1, intron E"
      exon      /note="PAI-1, intron F"
      intron     /note="PAI-1, intron G"
      exon      /note="PAI-1, intron H"
      intron     /note="plasmidogen activator-1"
      exon      /number=9

BASE COUNT      4974 a 4406 c 4386 g 3743 t
ORIGIN          131 bp upstream of SacI site; chromosome 7q21.3-q22.
```

```

Query Match      9.4%; Score 33.6; DB 10; Length 17509;
Best Local Similarity 48.0%; Pred. No. 2.5;
Matches 96; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY  95 aaacacttgaggagcagataactggccaaccatgactcagtgcttcttgaggccaaca 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  14196 AAAAATAAAGGAGCAGAAATCTGCTCAATGAGTAACAAGTCACCTACACTCCAAA 14137

QY  155 ggactctgagtcactcctgtgggtgaggtgggacaaagggggaatggtact 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  14136 TAACCCATGCACACATGTTCTCGGGAGGAGATGCCAGGCCAGGAGATTGGGCCACAT 14077

QY  215 gctgattacaacctgtgtgctgcctccctctcttctgagagggaagccatgc 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  14076 GATGGGGACATTCACCTCTGCCACCTGCAGCACCCCTGTACTGGGGAGGGGTGCCAGTG 14017

QY  275 ccaaaagtgtcacagccagg 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  14016 CCACAGTGACTCTGAGATG 13997
```

Search completed: June 24, 1999, 01:30:43  
Job time: 3031 sec

